CLAIR W. GOLDSMITH AND ROB KOLSTAD

a conversation about identity management

Clair W. Goldsmith is Associate Vice Chancellor and Chief Information Officer of the University of Texas System, and chair of the UT System Strategic Leadership Council.

CGoldsmith@utsystem.edu

Rob Kolstad is Editor of ;login:.

kolstad@usenix.org

THIS INTERVIEW WAS CONDUCTED BY Rob Kolstad with Dr. Clair W. Goldsmith on August 16, 2004.

RK: Please tell us a little about yourself and your workplace.

CG: I am the associate vice-chancellor and chief information officer of the University of Texas System. The system comprises 15 institutions, nine of which are general academic institutions and six of which are health institutions, that provide both health care and medical education. Many of the institutions are substantially involved in research as well.

The institutions employ about 81,000 and teach about 169,000 students, spanning the state of Texas in both rural and metropolitan regions.

RK: So you're associated with UT-Austin because you're located in Austin?

CG: I am part of the overall UT System Administration, an overarching organization. Our institutions are budgetarily and otherwise somewhat independent. As such, they have their own personnel offices, their own faculties, their own presidents, and so on. Their budgets are rolled up together and presented to the legislature by the folks in the office where I work. Individual institutions have roughly the same relationship to each other that Pontiac and Chevrolet do within General Motors.

RK: So you work more on overall strategy than tactical or day-to-day operations?

CG: Yes. It is largely a strategic job trying to figure out how to deal with the problems that confront the University, and then leverage the University's assets for the greater good of the greatest number.

My office has a focus that is guided by a governance structure with representatives from all the institutions. They have selected IT security, leveraging UT system-wide buying power to reduce costs, and the system-wide network as my current tasks.

RK: With almost 250,000 constituents, that must be a lot of purchasing power.

CG: Yes, it's a $7.8 billion operation. IT is, of course, a fraction of that, potentially approaching $1 billion.

RK: Let's talk about security. Any specific actions you're currently investigating?

CG: We do a lot of identification of products and services that will help our institutions perform their jobs more effectively and efficiently. We feel that one of the serious problems in the security arena is that we typically do not know precisely with whom we are dealing.

In order to deal with that, we postulate that we need a secure, scalable, standards-based interoperable iden-

tity-management infrastructure. Such a system is not simple! Being infrastructure, in many ways it is not terribly exciting to administrators who might rather purchase the latest and greatest payroll system, or maybe a course management system to meet the faculty's needs. Of course, these are valid wishes, but they all depend upon being able to manage identity.

RK: Is this identity management something like knowing exactly who is sending email (for spam management)? Or identity vis-à-vis cryptography in order to conceal communications? Or identity for purchases and contract signing? What kind of identities are you talking about?

CG: That is part of what we're talking about. We specifically define this arena as being composed of three aspects, one of which is "identity"—the name by which an individual or a service/entity is called.

RK: So you manage identities of non-human things as well?

CG: That's true; we manage the identities of various resources.

But let's simplify the discussion for the moment and talk about identity as it relates to carbon units. Those guys have things like names or social security numbers, both of which are valid identities—and you hear about those being stolen.

RK: Are we talking about the context of electronic communications?

CG: It really turns out that you can't stop there. Identity leads to the second aspect of this, which is authentication: "How certain are we that the identity we have been proffered relates to the particular individual who might be in front of us? By what credential or mechanism is that person authenticated so that the particular name is bound to that particular individual?"

This is different from signatures. It's more like going to the Notary Public and presenting your birth certificate, saying, "I am Joe Blow, and here are the papers from a trusted third party who swears that I am who I say I am." That is authentication.

The word can be used with respect to a document. The document is then "authenticated" because you somehow or other know something about the document, mainly if it is a typewritten document on a piece of paper. Whoever has read it attests in some way that it has not been altered and then signs it.

A digital signature provides assurance that the document is authentic—it hasn't been altered. Whether or not that digital signature is capable of being repudiated depends on whether or not that private key has been compromised and just how sure you are that the person who signed it is, in fact, the person they are supposed to be.

RK: I'm curious. How does one visit a Notary Public, carrying a birth certificate, for example, and then assert one's identity from some document created 15—or even 50—years ago? How can the Notary (or anyone) come to believe that the document relates to me?

CG: You've come to the word we need here: trust. At some point, you have to trust the document. In other words, the person who presented it to you has to have some ability to explain why they have the document and why they are trustworthy.

RK: Do we generally do a good job of this in USA society/culture right now? Passports, driver's licenses, etc.?

CG: Not particularly. Look at the 9/11 guys who went to the Virginia Department of Motor Vehicles and were issued driver's licenses because they had social security numbers. We do not do a good job of it.

RK: Perhaps they deserved a driver's license?

CG: I believe they did, but I recollect that their social security numbers were not valid.

RK: Do you believe that the sort of identity management you're talking about needs to be stronger than a 16-year-old going to the driver's license office and being issued a valid government ID card?

CG: I don't think any DMV ever expected their driver's licenses to be used, for example, for airplane boarding. I would like digital identity management to be at least that good as a lower bar.

RK: There seems to be a difference in impact for some of these identity issues. You or I trying to prove our age to buy a beer is a very different thing from the president of a company signing a contract that obligates his company to many different terms and exchanges millions or billions of dollars.

CG: This is the third aspect, which is "authorization." In other words, what is the identified and authenticated individual authorized to do? The more "valuable" the transactions, the more certain we must be of the identity authentication. That has to do with the level of "assurance," or how strongly we have authenticated the person with the digital credential and their "role."

RK: How do you perform identity management? Is it an embedded chip or perhaps a biometric identification? Or is it more like possession of a USB dongle with your key on it, perhaps enhanced by a password?

CG: Most of the identity management we perform is called "single factor," something you know, like a password, which turns out to be a "shared secret." The "dual factor," which is something people are moving to right now, requires both something one "has" and something one "knows," like a password and a key, e.g., a USB key, that sort of token. It contains the digital credential.

"Three factor" is something you know, something you have, and something you are, which might be a biometric like a fingerprint or a retina scan.

Higher-factor identity management is a longer-term goal right now.

RK: In the universe of almost 250,000 clients, this sounds pretty expensive.

CG: It is expensive. We have institutions that are moving to "two factor" management and have more than one institution using USB tokens. A couple more are using smart cards.

RK: Would I be doing this as a student to, for example, take an exam? Or is it more like something I'd use to cover privacy issues when I get my grades?

CG: Both of those. You might even use a token to get into and out of a dorm (in addition to a physical metal key).

RK: That raises privacy issues of monitoring locations of students, doesn't it?

CG: No, in universities just about everything is an educational record and can't be divulged without the signature of the student.

RK: But you're collecting such data?

CG: Yes.

RK: And it can't be divulged to law enforcement agencies?

CG: It can be subpoenaed through the standard subpoena process. It can't be just "coughed up."

Back to the infrastructure we're discussing, though, the digital identity, the digital credential is just one aspect of it. What do you do with it when you have it?

What you really want to do is go around and nail it to every telephone pole so that if you want to send me email, you can find my credential in the public/private crypto-key sense and encrypt the email and sign it. Then I can find your credential on some telephone pole to assure myself that the mail really is authentic. This is a little bit similar to the PGP scheme.

RK: What are the components of an identification management system that you require to put a scheme like this into practice?

CG: You need a directory service. We're looking at LDAP.

RK: What's the query to LDAP?

CG: It could be a number of different things.

RK: "I'd like a blond with blue eyes . . ."

CG: Absolutely. If she or he has agreed to release that information, we'll cough it up right away.

The second requirement is for a mechanism that can embed something like a defined (uniform) object into that directory so that those institutions who want to can share information about objects in that directory. In the case of higher education, we have chosen an object called "eduPerson," which is promulgated by EDUCAUSE and Internet2. It's a quasi-standards-based definition based on "inetOrgPerson."

RK: It's a name, address, phone number?

CG: Those things are part of inetOrgPerson; eduPerson has data that higher education might want. It has many, many attributes, like whether a person is a student or faculty member. We have fought bitterly over them for the last three years. It's quite extensible.

The eduPerson record is supposed to have the attributes that span the 3,000 higher education institutions in the USA. It might include a major, gender, age . . .

Then you can create "UTAustinPerson" with attributes that are the extension part that might be unique to UTAustin. This might be something like football tickets.

The next aspect of this is something developed by Internet2, and promulgated by Internet2 and EDUCAUSE, called "Shibboleth."

Shibboleth is a mechanism for institutions to share attributes about persons or entities in the LDAP directory. The sharing is, in the case of individuals, controllable by those individuals to some extent. In some sense, it has policy aspects to adjudicate requests.

Let me give you an example. Consider using JSTOR, the journal storage for past academic journals, a pan-university entity that stores documents. If I am a student at some university and want to access some document, I go to the "WAYF" ("Where Are You From") processor and select my institution. I contact JSTOR, who then asks my institution to get me to logon using my local credentials, login ID, and password—all from the LDAP directory.

RK: So JSTOR is somehow permitted to do all this?

CG: Yes, JSTOR is called a "resource provider" and there's a contractual arrangement with them. The next question it's going to ask is, "What is the person's role: faculty, staff, student, or maybe something more generic like 'member of community'?" Perhaps that's all the license requires. The answer might be specific (e.g., "student") or just "yes," the person is a member of the community, whatever is required to satisfy the contractual obligation. JSTOR wants to know if you meet the contractual requirements between the institution and JSTOR.

RK: Would Napster use this to authenticate downloads?

CG: This is exactly what is being used at Penn State.

RK: If UT students bought individual subscriptions, they wouldn't use this fancy management system, would they?

CG: They'd have nothing to do with it. This is for university business.

RK: Transcripts?

CG: Your transcript wouldn't be part of the system but access to the transcript might well be controlled by it.

RK: How many kilobytes in a typical student record?

CG: It's big. Maybe more than 100KB. The digital credential would be stored in the directory. Perhaps a faculty biography might be stored there. Generally, I view it more as containing items like a biography than as pointing to them. It might have group memberships like "member of English 101, section 1."

RK: Like a class schedule?

CG: No, it's memberships. A set of classes—and their members—might be derivable from the memberships, though. Strict privacy laws govern release of such data very strongly.

There's a goal here among some of the content providers and some of our business partners to reduce some of the bilateral contracts that exist.

RK: So there's a directory structure that doesn't sound so lightweight to me, and a permissions structure . . .

CG: You have an attributes release policy, which is fundamental to the system. I have a single PowerPoint slide that illustrates the Shibboleth mechanism; it takes 11 clicks to get through it.

RK: So Shibboleth is complicated?

CG: Yes, many protocol elements and policy implementers.

The identity portion is managed by the LDAP and Shibboleth pieces.

Exchanging things of value like contracts or credit card transactions or grant applications—anything that has to be signed and binds one or both parties to perform specific actions—is another matter.

The system must ascertain "authority" of the signee. Furthermore, the recipient has to trust or know with certainty that the signee has the proper authority to perform the transaction. How does the recipient trust such a thing in the case of a person they don't know?

RK: So we're talking about medical grants, for example, that move millions of dollars and require federal oversight for regulations more than we're talking about buying a CD with a credit card from a Web vendor?

CG: Yes. Let's talk about the mechanical process that has gone on in the past. Many people don't realize that a grant proposal to the National Institutes of Health includes the principal investigator's signature but also the signature of the University chief business officer or provost of research that is capable of binding the institution to the specific contractual rules and implicit laws and other context that exist between the institution and the granting agency.

RK: This is sounding very much more like legal issues than technical ones.

CG: Yes, it does. So NIH has on file, on paper, the signatures of every single person in the institutions of interest that is allowed to sign such documents. Historically, those signatures are compared manually to the signatures on the grant applications.

In the public key infrastructure (PKI) world, we have various levels of assurance. Some people want a "high assurance" digital certificate in order to sign big contracts. That technically is not the case here. The assurance here has to do with the authentication of the individual. In the federal scheme of things (and there is a mechanism and set of standards for using these tools), they have created a "bridge" in PKI terms. The bridge enables entities to establish transitive trust.

Say the Department of Education has a PKI system and the Department of Commerce also has a PKI system. How do you establish a linkage between them so that trust between them doesn't require (re-)issue certificates to members of the "other" department? What you want to do is put something in between the two PKI systems that is basically a set of policies that examines both entities and says, "For your assurance levels Red, White, and Blue, you'll require these properties of the other system," while the other system ends up with a similar statement, "For your assurance Bronze, Gold, and Platinum, you'll need . . ."—potentially a mapping between attributes of the two systems. This ultimately establishes transitive trust between the two agencies.

Interestingly, this mapping happens only at the bridge. Cross-certification happens only at the bridge. The agencies themselves continue to operate as before. That means if there's a federal bridge and NIH is cross-certified into the bridge and the UT system is cross-certified into the bridge, then a principal investigator and chief business officer at UT can digitally sign a contract and send it to NIH. NIH has an electronic process where that signature is verified by going back to source institution through the bridge, looking it up through the LDAP directory, and ensuring that the certificate is still valid and was valid at the time it was signed.

RK: These mapping institutions sound very complex, with an NxN matrix. No standards for this?

CG: We must do mappings at this point because we don't know yet what sort of standards are required. This might be more of an interim measure than a long-term solution.

RK: That would surely aid scaling. Why is this mechanism better than comparing signatures in a file cabinet?

CG: Good question. Some argue that, after 13 years of trying to implement this, perhaps this mechanism is too complex. There are a number of things, though, this provides for us.

If you're going to have an open Internet, which lots of us would like to see preserved, you're going to have to have secure transactions and trustable transactions on the network, even more than we trust credit card transactions now.

There are some types of transactions that occur at high frequency that can benefit from this type of system. Consider federal student loans. Right now, every loan is backed by a piece of paper in a file cabinet somewhere, and it's literally millions of loan applications every year.

RK: Computers have long been characterized by the accounting folks as backup for paper. Paper is "the real thing."

CG: Yes. If you'd like to get rid of paper, then you'll need an identity management system.

RK: How much will it cost to get rid of paper and save all that money?

CG: It's surely a chunk of money. But realize that right now we authenticate students, for example, hundreds of times during their tenure at an institution (e.g., library, health center, registration). It's distributed across a large number of organizations and costs a bit each time it's done. With identity management, it's conceivable that the hard part is done but once, and then authentication is simple and cheap.

There's also the legal requirements about discussing student data in email, for example. Any email that contains a social security number or student name or medical data must be encrypted. This infrastructure provides that capability, something we don't have now. It also reduces the potential for identity theft.

RK: And, of course, you can reject unsigned email and get rid of a lot of spam.

This is all hard to do, right?

CG: Not only is it hard to do, but it's all part of the infrastructure and, thus, is invisible to users and upper management. It's not as sexy as a new application, so this relates to what's really valuable to us as a society —the user doesn't want any of this right now.

RK: Would this system, in the future, displace Microsoft's Passport system or Amazon's one-click ordering by remembering your data and coordinating with those type of systems?

CG: That's certainly what Shibboleth in combination with PKI in fact can provide.

RK: The PKI has been talked about for years and years. PGP has attempted to implement some of it via machines at MIT and other places. How come we haven't seen more penetration and deployment of PKI? What are the challenges?

CG: Getting that infrastructure in place and deployed. I have one institution that's been up for about four years. They have three or four thousand certificates issued. They've been able to change passwords securely over the network. That's a big deal when your computer center is flooded with tens of thousands of gallons of water and everyone has to work at home, as happened there.

Dartmouth is issuing certificates to its incoming freshman this year

RK: All standardized and sharable?

CG: All X.509v3 and, in theory, sharable. Unfortunately, the contents of the certificates are not universally standardized. At UT, where we're trying to do the same profile across all the institutions, we're still encountering some conflict. Some institutions, for example, don't want to include a user's email address in the certificate, which is counterproductive.

RK: Do the Dartmouth freshmen get a physical token?

CG: I think the key goes on their laptop.

RK: So a stolen laptop is a stolen identity?

CG: No, you need the password, too.

RK: Ah, so it's both a stolen laptop and torture. Similar to some other systems. How does this all compare to extortion like, "Write me a check for $50,000 or I'll kill you"?

CG: About the same level.

RK: What non-technical constraints do you live under?

CG: The public isn't demanding it, and it's not on their conscience.

Logins and user IDs are proliferating. I have over a hundred. I don't put them into my browser; I write them to an encrypted file.

Of course, my browser remembers the non-critical passwords, like for the New York Times.

RK: And this identity management will be affordable?

CG: It's already affordable. LDAP is even available as "freeware."

We're doing it (though we're often doing it wrong), and we're doing it in many ways. Scalability will work.

RK: As you look forward and get the budget and start deploying this system, what do you see?

CG: There are detractors who says it's too complex and won't get off the ground. Until it's more widely deployed, it's always a possibility. My experience has been that I've been working for 13 years to get it out there. I haven't succeeded yet. But in my 13 years,

nothing has come close to displacing it or to being an alternative that is as effective as the system appears. A year from now, we'll have more progress, more Dartmouths, and more success stories.

Demonstrable savings will appear with more security and better individual privacy; we're slowly building momentum for this train.

RK: Thanks for sharing with us today!

Thanks to USENIX Supporting Members

Addison-Wesley/Prentice Hall PTR

Ajava Systems, Inc.

AMD

Aptitune Corporation

Asian Development Bank

Atos Origin BV

Delmar Learning

DoCoMo Communications Laboratories USA, Inc.

Electronic Frontier Foundation

Hewlett-Packard

Interhack

MacConnection

The Measurement Factory

Microsoft Research

Perfect Order

Portlock Software

Raytheon

Sun Microsystems, Inc.

Taos

UUNET Technologies, Inc.

Veritas Software

USENIX Supporting Member Benefits

One representative receives the benefits on behalf of the company. Join today! Send email to Catherine Allman at sales@usenix.org.

- Free subscription to *;login:*, the magazine of USENIX, both in print and online
- Online access to all Conference Proceedings, 1993–present
- All Conference Proceedings produced during the membership term can be downloaded to your institution's server, giving your students and staff full access to papers from our events
- Place one free ad in *;login:* ($1700 value) during the membership term
- Receive a 10% reduction in sponsorship and exhibit fees for USENIX-sponsored conferences, as well as premium placement on the exhibit floor
- Register up to ten staff at the member price for conferences during the membership term ($1100 value)
- Your click-through logo or company name on the USENIX Web site
- Acknowledgment in conference materials and *;login:*
- The right to vote in USENIX Association elections
- Discounts on technical registration fees for all USENIX-sponsored and co-sponsored events
- Discounts on purchasing Proceedings, CD-ROMs, and other USENIX publications
- Discounts on industry-related publications such as *Sys Admin, Linux Magazine*, and O'Reilly and No Starch Press books